Symposium on Reading Research

EVENT-RELATED BRAIN POTENTIALS, EYE MOVEMENTS, AND READING

Gary E. Raney and Keith Rayner
University of Massachusetts

Abstract—*This article compares studies that use event-related brain potential (ERP) and eye movement data to examine changes in reading behavior when a text is read twice. Although the types of information provided by these methodologies are different, both indicate that rereading a text facilitates many aspects of processing. ERPs provide a method for measuring comprehension and memory processes separately, while eye movements provide a continuous record of performance and allow changes in reading behavior to be localized to specific words. The results from these studies are compatible. However, converging evidence is not always found when different paradigms are contrasted, and diverging results can provide important information. To facilitate comparison across experiments, we suggest using a common set of materials for both paradigms. We conclude that comparing the results of research based on more than one paradigm provides a more complete understanding of the processes involved in reading.*

Reading is a complex skill based on a collection of mental processes. This complexity is reflected by the number of different research paradigms used to study reading. These methodologies are based on temporal measures (such as fixation duration and reading rate), measures of lexical access (such as naming and lexical decision time), and psychophysiological measures (such as event-related brain potentials, or ERPs). Each of these measures has made a contribution to the field, but each has strengths and weaknesses.

Since different paradigms provide unique sources of information, there is a need to integrate the findings of research based on different paradigms. The purpose of this article is to report how we have attempted to integrate the findings from studies using ERPs with studies using eye movements as measures of cognitive processing during reading. Before doing so, we first highlight the characteristics of each measure.

Address correspondence to Gary Raney, Department of Psychology, University of Illinois at Chicago, Box 4348, Chicago, IL 60680-4348.

EVENT-RELATED BRAIN POTENTIALS

As neurons in the brain are activated and inhibited, they produce extremely small changes in their surrounding electrical fields. When many neurons act in synchrony, their electrical fields combine to produce small voltage fluctuations which may be recorded from the scalp. These ongoing changes in electrical activity are commonly known as the electroencephalogram, or EEG. An ERP is formed from the portion of the EEG which represents the brain's response to a specific event in time. Because the ERP is embedded in the EEG signal, the ERP to a single stimulus can be very "noisy," or hard to discern. To increase the signal-to-noise ratio, responses to similar stimuli are averaged together to obtain an average ERP response (waveform), much as reaction times are averaged together to obtain a mean response time.

An ERP is described in terms of the latency and polarity of components in the waveform. The major peaks in the waveform are commonly referred to as components. For example, in Figure 1, the first component is labeled N100 because it is a negative-going peak with its maximum (negativity) around 100 ms after stimulus onset (note that negative is plotted up in Fig. 1). The next two components, labeled P200 and P300, are positive-going waves that peak around 200 ms and 300 ms after stimulus onset, respectively.

An important point is that different components are often sensitive to different types of cognitive processes. For example, if an auditory probe stimulus (i.e., a beep) is presented while a person is reading, the amplitudes of the N100 and P200 responses to the probe are altered by low-level attentional demands, whereas the amplitude of the P300 response to the same probe is altered by memory demands (Raney, 1993). (We expand on this example later.) In addition to changes in amplitude, the latency of components varies. For example, within a categorization task, P300 latency increases as the stimulus discrimination becomes more difficult (Gopher & Donchin, 1986). The appearance of later occurring components (after 200 ms) is dependent on task demands. For example, the N400 component (which is responsive to semantic aspects of a task) is not apparent in Figure 1. The task represented in Figure 1 (described later) did not involve the type of processing needed to produce an N400.

There are several reasons why ERPs are useful measures of cognitive processing. First, ERPs are an on-line measure, and the latency of changes in the waveform provides an indication of the timing of a process. Second, the differential sensitivity of ERP components to cognitive processes allows evaluation of potentially unique elements or stages of processing within a single task. Third, ERPs can be recorded nonintrusively, because no overt response to a stimulus is needed to produce an ERP.

EYE MOVEMENTS

When reading text, people alternate between fixations (the time when the eyes are stationary) and saccades (the time when the eyes are moving from one point to the next). Fixations typically last between 200 and 250 ms, and saccade length averages around eight char-

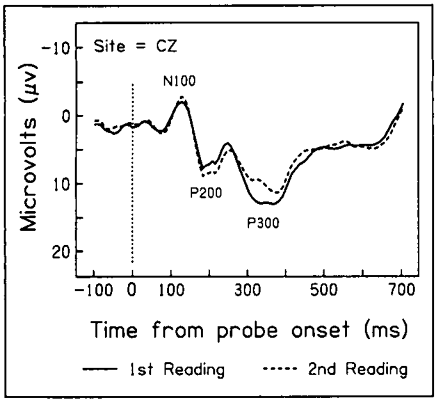

Fig. 1. ERP responses to secondary auditory probes during two readings of a text. Recording site is C_Z (midline center). ERP is for high "reading span" (reading ability) subjects when a speeded detection response to the probes was required (based on data from Raney, 1993, Experiment 1).

acter spaces. Although most saccades represent forward movement through a text, regressive movements are also made in order to reread prior material. Fixation duration, saccade length, and number of regressions vary considerably within subjects, and this variation is tied closely to processes associated with reading (Rayner & Pollatsek, 1987, 1989). For example, when a person is reading difficult text, fixation duration increases, saccade length decreases, and the frequency of regressions increases (Rayner, 1978).

The amount of time spent fixating a word is commonly reported as two different measures, first fixation duration and gaze duration. If a reader makes multiple fixations on a single word, first fixation duration includes only the very first fixation. Gaze duration includes all fixations made on a word before moving to another word. More than one fixation might be made on a word if the word is long or if additional time is needed to integrate the word's meaning into the preceding context. Though the two measures reflect similar processes (Rayner & Pollatsek, 1987), gaze duration may reflect additional integrative processes that are not reflected by first fixation duration.

Eye movements possess several strengths as a measure of reading behavior. First, monitoring eye movements produces a continuous, on-line record of reading performance. This record enables researchers to localize the cause of changes in performance to specific parts of a text. Second, and perhaps most important, eye movements are a natural part of reading, so no additional or unusual task demands are placed on a subject. Third, multiple aspects of eye movements may be analyzed (e.g., fixation duration, saccade length, and frequency of regressions), which provides a window into different elements of the reading process.

Although ERP and eye movement recordings represent substantially different techniques, in many ways their advantages are similar. Both are on-line measures of performance, neither requires a response which is not part of natural reading, and each is a "data rich" measure which may be broken down into subcomponents that reflect different processing demands.

The ideal situation for comparing ERPs and eye movements would be to record both measures simultaneously. Two limitations on ERP recording currently make this very difficult, however. As the eyes move during reading, the ocular muscles produce substantial interference in the ERP. To reduce this problem, ERP studies tend to present sentences one word at a time in a central location. This procedure eliminates the need for eye movements, but makes the reading situation less natural. However, techniques have been devised to "subtract" out this interference, which may eliminate this problem in the future (Berg & Scherg, 1991; Gratton, Coles, & Donchin, 1983). Even if this problem is solved, a potentially more serious problem is component overlap. The components of interest in language studies usually occur from 200 to 600 ms after stimulus onset. Since the average fixation duration on a word is 200 to 250 ms, components of the ERP response to a word which occur later than 250 ms could overlap with the ERP response to the next word. Complete separation of overlapping responses might not be possible.

Because recording ERPs and eye movements concurrently is problematic, the approach we have taken is to compare the results from separate studies based on the two paradigms. By investigating similar issues using the two techniques, we have found converging results regarding some issues, and diverging results regarding others. This pattern of convergence and divergence has served to strengthen some findings while clarifying others. The following discussion provides an example of how we are contrasting ERP and eye movement studies.

COMPARING ERPS AND EYE MOVEMENTS

The prevailing method for manipulating processing demands is to contrast reading performance across two types of stimuli. In this approach, differences in stimuli produce changes in mental processes. An alternative approach is to manipulate the situation while holding the stimuli constant. An example of this is rereading. When an individual reads a passage of text twice, there are processing changes associated with the second reading. Although rereading is a common behavior, especially in educational settings, rereading has not been a major topic of reading research. We feel that repeated exposure to a text provides an opportunity to vary naturally the processes associated with reading and will shed light on such processes. In the following discussion, we describe how we have investigated rereading using ERPs and eye movements.

Prior research has shown that when a text is read twice, it is read faster during the second reading (Hyona & Niemi, 1990). This finding implies reduced cognitive load (processing difficulty) during the second reading. An important question is, what factors lead to reduced load? The first study we describe used ERPs to examine this question.

In a series of experiments using a secondary task paradigm to measure cognitive load during reading (Raney, 1993), subjects read short passages (the primary task) while listening to a series of secondary auditory probes (i.e., beeps). ERP responses to the secondary probes served as the index of cognitive load in the primary task. Two ERP components were examined: the N1-P2 (defined as the difference in amplitude between the N100 and P200) and the P300. Prior work had shown that, on the one hand, N1-P2 amplitude reflects low-level attentional demands in the primary task, such as attentional focus and stimulus evaluation (Hillyard & Kutas, 1983). On the other hand, P300 amplitude had been shown to be sensitive to high-level attentional demands of the primary task, such as memory updating (Gopher & Donchin, 1986). For both the N1-P2 and the P300, an increase in the amplitude of responses to secondary probes indicates reduced cognitive demands by the primary task. If a second reading demands fewer resources for both lower and higher level processes, both the N1-P2 and the P300 should increase in amplitude during the second reading. In other words, if reading becomes easier during the second reading, subjects will notice the probes more and, consequently, should have larger amplitude N1-P2 and P300 responses to the probes.

In the experiments, the subjects' task was to remember as much of the passages as possible for later recall. Each text was displayed one word at a time in a common location at a rate of 120 words per min. Two results are relevant here. First, the amplitude of the N1-P2 increased during the second reading, which indicates less cognitive load. Second, P300 amplitude decreased during the second reading, which indicates greater cognitive load.[1] The change in N1-P2 and P300 amplitudes can be seen in Figure 1.

These apparently conflicting results were interpreted as reflecting task demands and the sensitivity of different ERP components to different processes. The N1-P2 was described as reflecting lower level demands, such as initial perceptual and comprehension processes. The P300 was described as reflecting higher level demands, such as memory updating. Lower level demands decreased during the second reading (i.e., the texts were easier to read and comprehend), whereas higher level memory demands increased during the second reading (i.e., more material was held in memory during the second reading). The increase in memory demands resulted from task instructions, which emphasized memory of the material. These results indicate that more than one factor contributes to processing load during reading and that these factors may be separable. The resource demands of each factor appear to vary depending on task demands.

This study demonstrates that lower level processes are facilitated during a second reading. What aspects of reading behavior reflect this facilitation? In the following study, we attempted to determine the aspects of reading performance that are facilitated (speeded) during a second reading.

We (Raney & Rayner, 1991) had subjects read short texts, each twice in succession, while we monitored their eye movements. Our goal was to determine whether a single factor, such as reduced fixation duration or increased saccade length, or multiple factors lead to increased reading speed during a second reading. In addition, we examined whether the standard word frequency effect (high-frequency words are read faster than low-frequency words) is modulated by a second reading. Of specific interest was whether changes in reading performance during a second reading would apply equally to low- and high-frequency words. To examine this second issue, a target word of low or high frequency was embedded in each text, and fixation times for the targets were compared. To control for meaning changes, targets were paired: Each pair consisted of two related words or synonyms, one of low frequency (e.g., *blossom*) and one of high frequency (e.g., *flower*).

Table 1. *Mean first fixation durations (ms) and gaze durations (ms) during two readings of a text (from Raney & Rayner, 1991)*

Target word frequency	First reading	Second reading
	First fixation duration	
Low frequency	287	271
High frequency	254	246
	Gaze duration	
Low frequency	343	314
High frequency	279	254

Reading speed did increase during the second reading, and this increase was reflected by multiple measures of performance. During the second reading, forward fixation duration was shorter, the number of forward fixations decreased, and forward saccade length increased. The number of regressive fixations was also less during the second reading, but regressive fixation duration and regressive saccade length did not change.

Regarding the frequency manipulation, we found that low-frequency words were read more slowly than high-frequency words during both readings, which replicated the standard word frequency effect (see Rayner & Duffy, 1986). Both low- and high-frequency words were read faster during the second reading and, interestingly, the proportional decreases in fixation times were roughly equal. That is, there was no interaction between word frequency and repetition for either first fixation duration or gaze duration for the target words (see Table 1).

As does the ERP study of rereading, this study supports a multicomponent description of reading. Rereading led to a change in several performance measures, including fixation duration, number of fixations, and saccade length. Furthermore, repetition effects did not inter-

1. Results are presented here only for recording site C_Z (midline center). ERPs were also recorded from two frontal and two parietal sites. The pattern of results for the N1-P2 varied depending on the recording site, whereas P300 results were consistent across sites. In addition, results based on the N1-P2, but not on the P300, varied based on each subject's reading ability. Half of the subjects made an overt detection response to the probes and half made no response to the probes. N1-P2 amplitude was unaffected by response condition, but P300 amplitude was larger in the response than in the no-response condition.

act with frequency effects, which implies independent (Sternberg, 1969) or modular (Fodor, 1983) stages of processing.

Taken together, the ERP and eye movement studies support the conclusion that changes in rereading behavior are reflected by multiple factors. The ERP study (Raney, 1993) showed that lower level perceptual and comprehension demands are independent of higher level memory demands and the resources required by these processes change from a first to a second reading. The eye movement study (Raney & Rayner, 1991) showed that more than one measure of performance reflects facilitation (i.e., increased reading speed) during a second reading and that this facilitation is not limited to either low- or high-frequency words. Both the ERP and the eye movement studies support a componential description of reading.

The studies just described provide different, but compatible, types of information about rereading. However, converging results will not always be found. In our own research, we have also found diverging results when comparing ERP and eye movement studies. A brief example illustrates the usefulness of comparing studies even when diverging results are found.

Morris (1992) examined how a sentence context facilitates reading of words. Subjects read sentences such as "The barber trimmed the mustache in the morning" while their eye movements were monitored. In some sentences, the subject noun (*barber*) and the verb (*trimmed*) were mildly related to a target word (*mustache*). In control sentences, the subject noun and verb were replaced by neutral words (e.g., *man* and *saw*). Morris found that reading the target words was facilitated (speeded) in the original sentences compared with the control sentences. In addition, if either the subject noun or the verb alone was replaced by a neutral word, no facilitation was found. This result suggests that, individually, the subject noun and the verb did not possess a strong enough relation to the target to facilitate its processing. Raney, Fischler, and Hardonk (1992) conducted an analogous experiment, but recorded ERP responses to the target words, and found facilitation (based on N400 amplitude) for sentences which contained both the original subject noun and the original verb as well as for sentences in which only one of those words was replaced. Thus, ERPs reflected priming when fixation times did not.

Interpreting these different results is beyond the scope of this article. The point we want to make is that two paradigms do not always provide the same answers to a question. Results that differ across paradigms serve to point out an area of study needing clarification. In the prior example, eye movements and ERPs might provide opposing data because they reflect different aspects of processing, because one measure is more sensitive than the other, or because the methodologies are not identical. Regardless of whether eye movements and ERPs reflect similar processes, both measures provide information which will enhance understanding of the role of context in sentence processing. Diverging results might provide as much useful information as converging results.

SUMMARY AND CONCLUSIONS

The studies we described demonstrate the usefulness of contrasting the results of research from different paradigms. As is true for ERPs and eye movements, directly combining two measures is not always feasible. Our approach, which is not limited to ERPs and eye movements, has been to compare studies that address similar issues. The most fruitful comparisons can be made when studies use common materials. When this is done, converging results become stronger, and diverging results have a better chance of being understood; when one paradigm is compared with another, both paradigms benefit.

Acknowledgments—This research was supported by National Institute of Mental Health Grant MH16745 awarded to the Department of Psychology and by National Institutes of Health Grant HD26765 and National Science Foundation Grant DBS-9121375 awarded to Keith Rayner. We wish to thank Ira Fischler for his helpful comments on this manuscript and discussions of this material.

REFERENCES

Berg, P., & Scherg, M. (1991). Dipole models of eye movements and blinks. *Electroencephalography and Clinical Neurophysiology, 79,* 36–44.

Fodor, J.A. (1983). *Modularity of mind.* Cambridge, MA: MIT Press.

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance: Vol. 2. Cognitive processes and performance* (pp. 41.1–41.49). New York: Wiley.

Gratton, G., Coles, M.G.H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55,* 468–484.

Hillyard, S.A., & Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual Review of Psychology, 34,* 33–61.

Hyona, J., & Niemi, P. (1990). Eye movements during repeated reading of a text. *Acta Psychologica, 73,* 259–280.

Morris, R.K. (1992). Sentence context effects on lexical access. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 317–332). New York: Springer-Verlag.

Raney, G.E. (1993). Monitoring changes in cognitive load during reading: An event-related brain potential and reaction time analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 51–69.

Raney, G.E., Fischler, I., & Hardonk, M. (1992, November). *ERP evidence of lexical and message level priming in sentences.* Paper presented at the meeting of the Psychonomic Society, St. Louis, MO.

Raney, G.E., & Rayner, K. (1991, November). *The effects of word frequency during two readings of a text.* Paper presented at the meeting of the Psychonomic Society, San Francisco.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85,* 618–660.

Rayner, K., & Duffy, S. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14,* 191–201.

Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 327–362). London: Erlbaum.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. In W.G. Koster (Ed.), Attention and performance II. *Acta Psychologica, 30,* 276–315.